

基础研究

DOI:10.13406/j.cnki.cyxh.002189

基于生物信息学技术挖掘番茄红素 对前列腺癌细胞的作用靶点

肖仕全¹, 王 利¹, 刘太行², 俞焱阳¹, 付利娟²

(1. 重庆医科大学附属第三医院生殖医学中心; 2. 重庆医科大学中医药学院中药综合教研室, 重庆 400016)

【摘要】目的:前列腺癌是泌尿系统中发病率最高的肿瘤,番茄和其他含番茄红素的食物对预防前列腺癌和减缓其发展进程有效,但是番茄红素作用于前列腺癌的机制尚不明确。本研究利用生物信息学方法深入挖掘番茄红素作用于前列腺癌的潜在靶点,有利于进一步阐明其在前列腺保护中的分子机制以及潜在的其它应用。**方法:**从 GEO 数据库检索并下载番茄红素补充和安慰剂对照的列腺癌患者前后的癌组织表达组基因芯片数据集,并利用加权基因共表达网络分析,构建基因共表达模块,并对性状高关联模块进行 GO 及 KEGG 聚类分析和节点基因筛选。**结果:**利用数据集 GSE27140 中 114 例样本,共构建了 16 个共表达模块,其中 2 个模块与性状关联度较高,参与上皮细胞-间充质细胞转变、姐妹染色单体分离、细胞衰老等信号通路。进一步对模块基因连通性分析发现 2 个节点基因。**结论:**本研究筛选出番茄红素治疗前列腺癌的潜在作用靶点,为阐明番茄红素的作用机制提供了有意义的线索。

【关键词】前列腺癌;番茄红素;加权基因共表达网络分析;生物信息

【中图分类号】R966

【文献标志码】A

【收稿日期】2019-05-15

Mining targets of lycopene on prostate cancer cells based on bioinformatics technique

Xiao Shiquan¹, Wang Li¹, Liu Taihang², Yu Zhiyang¹, Fu Lijuan²

(1. Center of Reproduction, the Third Affiliated Hospital, Chongqing Medical University;

2. Teaching and Research Section of Traditional Chinese Medicine,

College of Traditional Chinese Medicine, Chongqing Medical University)

【Abstract】Objective: Prostate cancer is the most common tumor in the urinary system. Tomato and other foods containing lycopene are effective in preventing prostate cancer and slowing down its development, but the mechanism of action of lycopene on prostate cancer is not clear. In this study, bioinformatics methods were used to explore the potential targets of lycopene acting on prostate cancer, which is helpful to further elucidate the molecular mechanism of action of lycopene in prostate protection and other potential applications. **Methods:** The microarray dataset of gene expression from prostate cancer tissues of patients supplemented with lycopene and placebo as control before and after treatment were searched for and downloaded from gene expression omnibus (GEO). The gene co-expression module was constructed by weighted gene co-expression network analysis. GO and KEGG cluster analyses and hub gene screening were performed on modules highly associated with traits. **Results:** A total of 16 co-expression modules were constructed from 114 samples in the dataset GSE27140. Two of the co-expression modules were highly associated with traits, which involved in epithelial-mesenchymal cell transformation, sister chromatid separation, cell senescence and other signaling pathways. Two hub genes were identified through further analysis of module gene connectivity. **Conclusion:** This study has screened out the potential targets of lycopene in the treatment of prostate cancer, which provides a meaningful clue for elucidating the mechanism of action of lycopene.

【Key words】prostate cancer; lycopene; weighted correlation network analysis; bioinformatics

作者介绍: 肖仕全, Email: 83948936@qq.com,

研究方向: 生殖药理学。

通信作者: 付利娟, Email: fulijuan@cqmu.edu.cn。

基金项目: 重庆市委基础与前沿研究计划资助项目 (编号:

cstc2016jcyjA0247)。

优先出版: <http://kns.cnki.net/kcms/detail/50.1046.R.20190530.0859.002.html>

(2019-05-30)

前列腺癌(prostate cancer, PCa)是男性生殖系统常见的恶性肿瘤,占美国男性新发癌症的首位,约占 20%。在中国,PCa 自 2008 年起成为泌尿系统中发病率最高的肿瘤。临床上 PCa 常通过手术或药物进行去除雄激素的内分泌治疗,有较好疗效,但往往在 18~24 个月后发展为去势抵抗性前列腺癌,最终导致患者死亡。PCa 的发病机制主要涉及遗传和环境因素,但仍未完全阐明。生活方式和饮食习惯对 PCa 的发生发展有着至关重要的影响。

番茄红素(lycopene)又称 ψ -胡萝卜素,属于异戊二烯类化合物,是水果和蔬菜红颜色的主要组成成分,是目前已知的抗氧化性最强的类胡萝卜素^[1]。服用番茄红素能明显降低罹患 PCa 的危险。有队列研究和相关调查发现,番茄红素的摄入能降低 PCa 的发生风险,特别是降低致死性前列腺癌的发生,不但能降低患病危险性而且能减缓前列腺癌转移的进程^[2-4]。番茄红素对前列腺癌作用的具体分子机制至今未明。利用相关组学以及生物信息学数据挖掘方法,探究番茄红素对 PCa 细胞的作用和途径,有利于进一步阐明其在前列腺保护中的分子机制以及潜在的其他应用。

1 材料与方法

1.1 番茄红素补充处理前列腺癌患者样本的微阵列数据基因表达分析

从美国国立生物技术信息中心(NCBI,网址:https://www.ncbi.nlm.nih.gov/)数据库的 GEO 子库中,使用 GEOquery^[5]工具下载 GSE27140^[6]数据集的基因表达矩阵和样本分组信息,筛选番茄红素补充组和安慰剂对照组共 114 例样本用于分析,样品分组信息详见表 1。

表 1 GSE27140 数据集中 114 例样品分组信息(例)

组别	处理后 0 个月	处理后 3 个月
安慰剂组	28	28
番茄红素补充组	29	29

微阵列注释信息 GPL13136 用于将探针与相应的基因匹配,对于多个探针对应于同一基因数据,计算该基因平均表达值。WGCNA 算法^[7]用于评估基因表达值。此外,使用 R (3.5.1)语言中 Hclust 工具执行聚类分析,筛选具有适当阈值的样品纳入后续分析。

1.2 共表达模块构建分析

基因共表达模块的识别和构建主要基于 R 软件包中加权基因共表达网络分析(weighted correlation network analysis, WGCNA)算法。首先利用梯度法用于测试不同软阈值下(软阈值范围 1~30)不同模块的独立性和平均连通度,当独立程度为 0.85 时,确定最适当的软阈值。确定软阈值后,继续基于 WGCNA 算法构造模块。为了确保结果的可靠性,将最小基因数设定为 40,进而提取每个模块中相应基因信息。最后使用热图工具包绘制热图以分析交互的强度。

1.3 构建模块-性状关系

基于模块 Eigengene(基因和样本构成的矩阵)和表型之间的相关性,估算模块-性状关联度,以鉴定与表型高度相关的表达组(模块)。进一步计算基因明显性(gene significance, GS)和模块身份(module membership, MM)的相关性。GS 为表达谱与每种性状之间相关性的绝对值;MM 为给定基因表达谱与给定模块的 Eigengene 的相关性^[8]。R 软件包 Pheatmap 用于基因表达热图的绘制。

1.4 共表达模块的功能富集分析及节点基因筛选

利用 R 软件包 ClusterProfiler,对性状高关联模块中基因进行生物学和信号通路聚类分析^[9]。基因聚类(gene ontology, GO)包括生物学进程(biological process, BP)、细胞组分(cellular components, CC)、分子功能(molecular function, MF)^[10]。KEGG 聚类(kyoto encyclopedia of genes and genomes)是将靶标基因映射到信号通路的富集分析^[11]。将 $P < 0.05$ 的聚类结果定义为明显聚类。通过 Cytoscape 软件对模块可视化,通过与其他模块基因的连接强度相加并将该数除以最大模内连通性来计算每个基因的模块内连接性,将具有最大模块内连接的基因确定为节点基因^[12]。

2 结果

2.1 番茄红素补充前列腺癌患者样本的共表达模块的构建

将 114 例番茄红素和安慰剂处理前列腺癌样品后测序得到的 9 523 个基因表达值通过 WGCNA 包装工具构建共表达模块;进一步通过 Hclust 工具对这些样品进行聚类分析,结果显示所有样本无异常情况均通过了阈值筛选(图 1)。软阈值作为共表达模块构建中最关键的参数之一,主要影响共表达模块的独立性和平均连通度,结果显示当软阈值等于 6 时,独立度达到 0.85 并且平均连通度较高(图 2)。基于软阈值 6 用于构建共表达模块,模块分别以不同颜色标注,并通过模块中包含的基因数量由多到少排序,结果显示共得到 16 个不同的基因共表达模块(图 3)。为了进一步评估每个模块之间的相似性以及向异性,根据基因间表达量进行聚类获得各模块间的相关性(图 4)。



图2 软阈值功率的网络拓扑分析

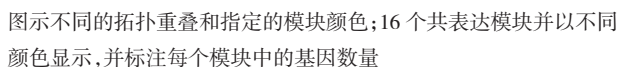


图 3 基因聚类树状图

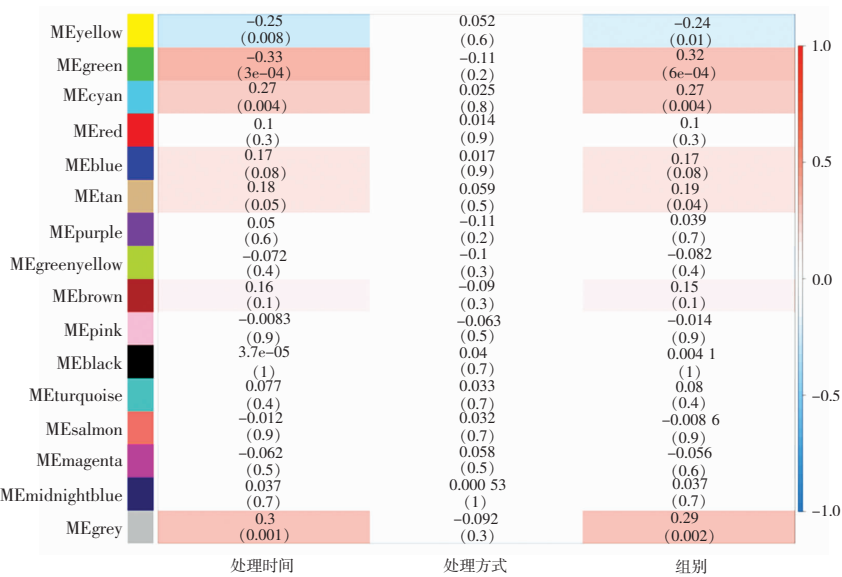


图 4 Eigengene 分析模块间相关性

2.2 基因共表达模块与表型的关联分析

为了确定 16 个已鉴定的表达模块中是否有模块与番茄红素处理相关,本研究对模块 Eigengene 和样本表型之间的相关系数进行了计算,结果表明其中 2 个模块(绿色、青色)与番茄红素处理明显相关(图 5)。进一步计算基因表达和模

块及性状之间的关联度,结果进一步确认了 2 模块中基因表达与番茄红素处理具有明显相关(图 6)。将青色模块 77 个基因与绿色模块中 511 个基因在 114 例样本中的表达情况进行热图分析,直观展示模块基因在不同样本的表达情况(图 7)。



y轴表示不同模块 eigengene;x 轴标识不同表型;取样时间:处理后 0/3 个月;处理方式:包括番茄红素补充组/安慰剂、对照组;组别:取材时间 + 处理;单元格显示相应的相关性和 P 值,相关性强度以颜色强度编码

图 5 模块 - 表型关联图

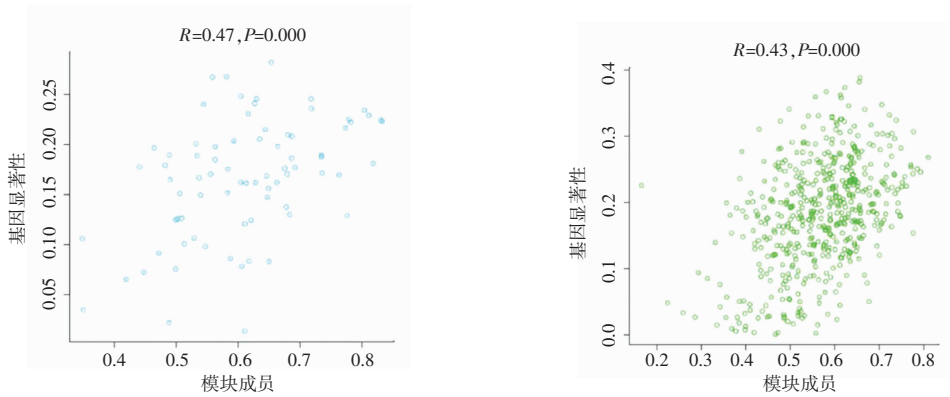


图 6 模块与基因表达的相关性散点图

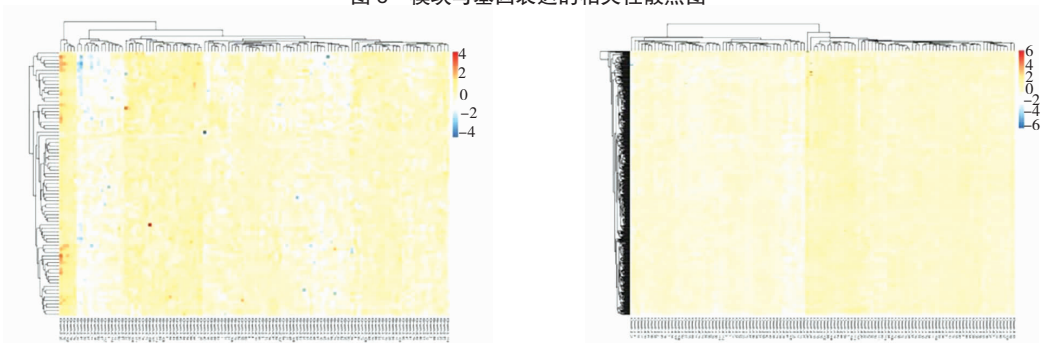


图 7 模块内基因表达热图

2.3 青、绿 2 模块基因的功能富集分析

进一步将 2 个模块中的基因进行 GO 及 KEGG 富集分析显示,2 个模块的生物学过程存在明显差异。青色模块中的基因主要富集在 RNA 剪接、上皮细胞-间充质细胞转变的正调节等通路(图 8);绿色模块的基因主要聚类于姐妹染色单体分离、细胞衰老等通路(图 9)。

2.4 模块可视化及节点基因筛选

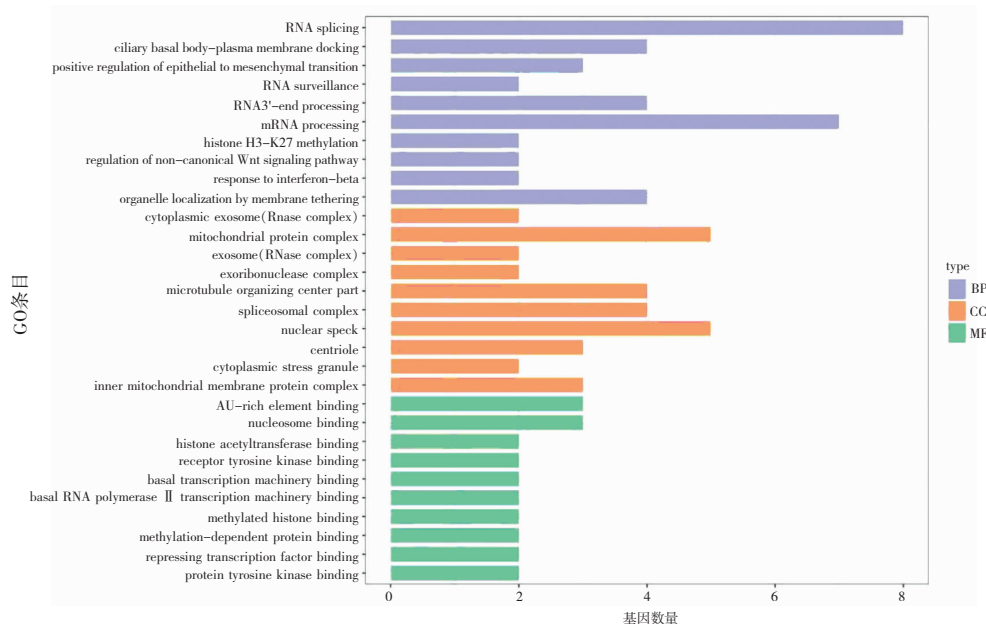
使用 Cytoscape 软件构建模块并计算模块内连接性,具有高模块内连接性的基因被认为是模块内节点基因(图 10)。青色模块节点基因为 PRPF4B,绿色模块节点基因为 SPPL2A。

3 讨论

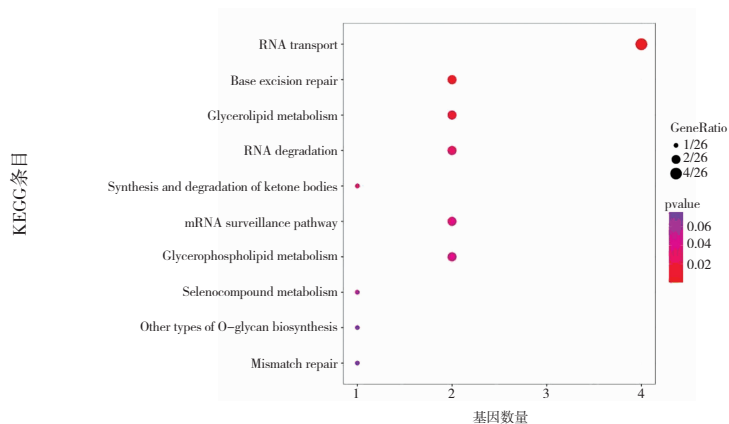
番茄红素对前列腺保护作用的主要机制包括清除自由基、诱导 II 相药物代谢酶 GSTP1,通过一系列的达到解毒目的,保护 DNA 及蛋白质免

受损伤^[13];抑制 PCa 癌细胞的增殖和促进凋亡诱导,将细胞周期阻滞于 G0/G1 期^[14];降低周期相关蛋白 Cyclin D1 的表达^[14]。番茄红素亦能调节 IGFs 信号通路,通过抑制 IGF-I 导致 IGF-IR 活化被抑制^[15],且增加 IGFBP3 的表达和分泌从而抑制前列腺癌细胞生长,增强多西他赛的抗前列腺癌作用^[15]。此外,番茄红素具有调控 DNA 甲基化的表观遗传治疗作用^[16]。番茄红素通过发挥抑制 DNMTs 功能,使抑癌基因高甲基化的启动子区 CpG 岛去甲基化,从而恢复这些抑癌节点基因的表达^[17]。

应用生物信息学研究药物作用靶点、分析机体的基因差异和各种功能差异,为药物的临床前研究提供大量相关资料,有利于提高药物应用的安全性和可靠性,亦能够发现机体对药物反应的靶点基因,解释药物对不同个体作用的差异和特点。本研究通过 WGCNA 分析,筛选番茄红素作用于前列腺癌的

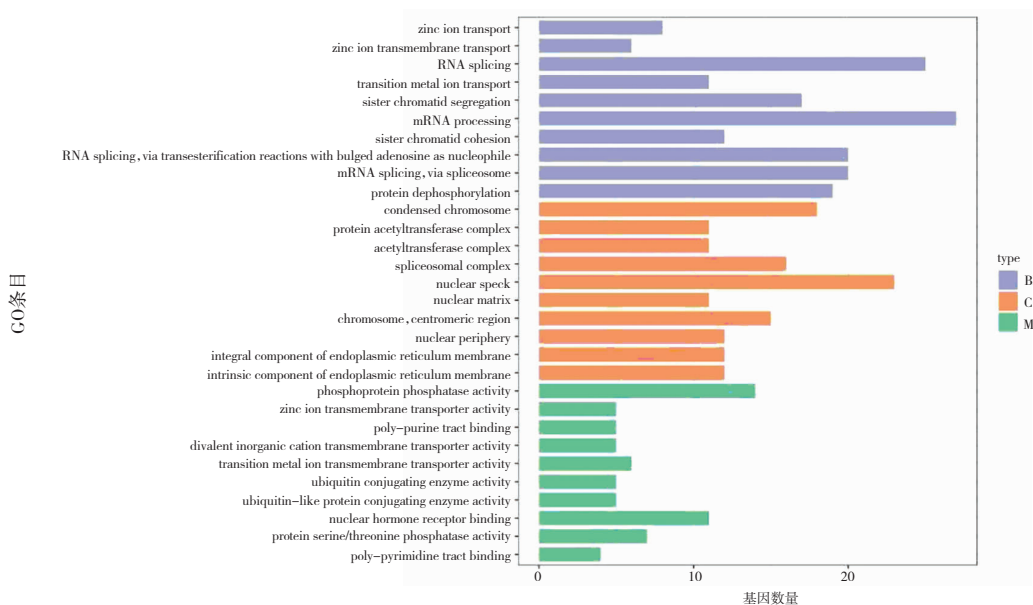


A. 青色模块中 77 个基因的 GO 聚类分析

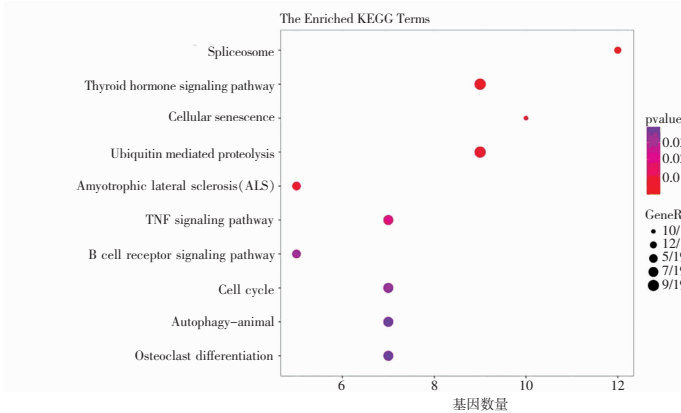


B. 青色模块中 77 个基因的 KEGG 通路聚类分析

图 8 青色模块中基因功能富集分析

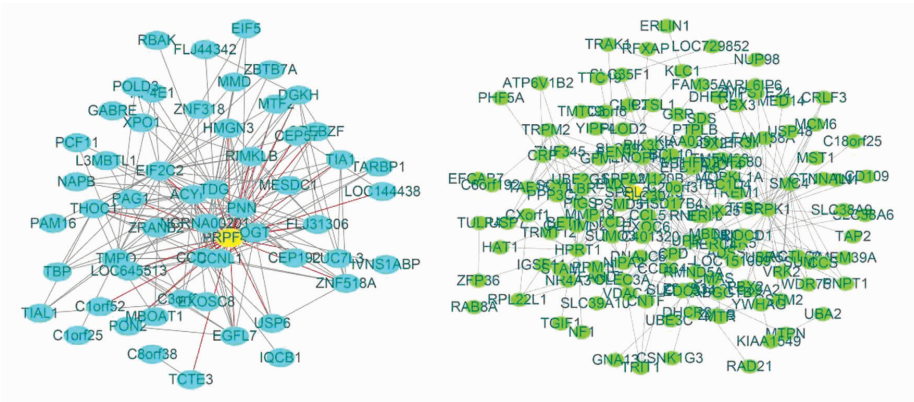


A. 绿色模块中 511 个基因的 GO 聚类分析



B. 绿色模块中 511 个基因的 KEGG 通路聚类分析

图 9 绿色模块中基因功能富集分析



A. 青色模块可视化, 模块中的节点基因以黄色标识

B. 绿色模块可视化

图 10 模块可视化及节点基因筛选

潜在靶点,确定其中 2 个基因共表达模块与组别具有相关性,并进一步计算 2 模块中的节点基因分别是 mRNA 前体加工因子 4B(PRP4B)和基因编码信号肽酶样 2A(SPPL2A)可能为番茄红素作用于

PCa 细胞的靶标基因。

PRPF4B 激酶属于类 CDK 激酶家族成员,参与 mRNA 前体剪切和信号转导。PRPF4B 的过表达可以通过激活细胞存活信号蛋白 NF- κ B、ERK 和 c-

MYC抑制白藜芦醇诱导的结肠癌细胞凋亡,能够通过减少 RhoA 蛋白的活性,实现对肌动蛋白细胞骨架的重排以及促进细胞由上皮型向间质型转化,这均显示 PRPF4B 与细胞的迁移和侵袭能力相关^[18]。

SPPL2A 基因编码天冬氨酸蛋白酶家族的一个成员,是位于跨膜区的 2 个保守催化基序的跨膜蛋白,也是信号肽酶样蛋白酶(signal peptide peptidase-like protease)家族的成员,其功能包括蛋白质同聚活性和天冬氨酸型内肽酶活性。可以对肿瘤坏死因子(tumor necrosis factor, TNF)等进行加工催化 TNF- α 锚定片段的膜内断裂,促进胞内结构域向细胞核信号的释放,还负责膜内切割 Fas 抗原配体,促进胞内 FasL 结构域的释放,在调节先天和适应性免疫方面发挥作用^[19]。近年来,SPPL2A 的不同家族成员在机械特性和底物要求方面的明显差异也已被解开,但是对其蛋白酶的病理生理功能的了解还远未完成。但是,基于对小鼠的研究已经有了很大的突破,SPPL 蛋白酶的抑制被认为是一种新的治疗手段^[20]。本研究生物信息学挖掘结果显示,SPPL2A 亦可能是 PCa 作用的潜在靶标之一,但是目前在 PCa 中的作用仍有待深入研究。

番茄红素作用于 PCa 细胞的靶标可能远不止于对一些靶标基因的表达作用,其对相关信号通路的激活亦可能是其作用机制的一部分,如对 TNF 信号通路的作用,该信号通路已被广泛证实参与了 PCa 的发生和发展^[21]。同样的,上述富集的潜在靶标基因 SPPL2A 亦是 TNF 信号通路中的一员。此外,如泛素化、细胞周期以及自噬等信号通路均参与到番茄红素作用于 PCa 细胞的效应中,值得深入探索。

参 考 文 献

- [1] Schwarz S, Obermuller-Jevic UC, Hellmis E, et al. Lycopene inhibits disease progression in patients with benign prostate hyperplasia [J]. The Journal of nutrition, 2008, 138(1): 49-53.
- [2] Giovannucci E, Ascherio A, Rimm EB, et al. Intake of carotenoids and retinol in relation to risk of prostate cancer [J]. Journal of the National Cancer Institute, 1995, 87(23): 1767-1776.
- [3] Zu K, Mucci L, Rosner BA, et al. Dietary lycopene, angiogenesis, and prostate cancer: a prospective study in the prostate-specific antigen era [J]. Journal of the National Cancer Institute, 2014, 106(2): djt430.
- [4] Key TJ, Appleby PN, Allen NE, et al. Plasma carotenoids, retinol, and tocopherols and the risk of prostate cancer in the European Prospective Investigation into Cancer and Nutrition study [J]. The American Journal of Clinical Nutrition, 2007, 86(3): 672-681.
- [5] Davis S, Meltzer PS. GEOquery: a bridge between the Gene Expression Omnibus (GEO) and BioConductor [J]. Bioinformatics, 2007, 23(14): 1846-1847.
- [6] Magbanua MJ, Roy R, Sosa EV, et al. Gene expression and biological pathways in tissue of men with prostate cancer in a randomized clinical trial of lycopene and fish oil supplementation [J]. PLoS One, 2011, 6(9): e24004.
- [7] Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis [J]. BMC bioinformatics, 2008, 9: 559.
- [8] Shi K, Bing ZT, Cao GQ, et al. Identify the signature genes for diagnose of uveal melanoma by weight gene co-expression network analysis [J]. International Journal of Ophthalmology, 2015, 8(2): 269-274.
- [9] Yu G, Wang LG, Han Y, et al. clusterProfiler: an R package for comparing biological themes among gene clusters [J]. OMICS, 2012, 16(5): 284-287.
- [10] Ashburner M, Ball CA, Blake JA, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium [J]. Nature Genetics, 2000, 25(1): 25-29.
- [11] Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes [J]. Nucleic Acids Research, 2000, 28(1): 27-30.
- [12] Tan N, Chung MK, Smith JD, et al. Weighted gene coexpression network analysis of human left atrial tissue identifies gene modules associated with atrial fibrillation [J]. Circulation Cardiovascular Genetics, 2013, 6(4): 362-371.
- [13] Sheehan D, Meade G, Foley VM, et al. Structure, function and evolution of glutathione transferases: implications for classification of non-mammalian members of an ancient enzyme superfamily [J]. The Biochemical journal, 2001, 360(Pt1): 1-16.
- [14] Palozza P, Parrone N, Simone RE, et al. Lycopene in atherosclerosis prevention: an integrated scheme of the potential mechanisms of action from cell culture studies [J]. Archives of Biochemistry and Biophysics, 2010, 504(1): 26-33.
- [15] Tang Y, Parmakhtiar B, Simoneau AR, et al. Lycopene enhances docetaxel's effect in castration-resistant prostate cancer associated with insulin-like growth factor I receptor levels [J]. Neoplasia, 2011, 13(2): 108-119.
- [16] Huang J, Plass C, Gerhauser C. Cancer chemoprevention by targeting the epigenome [J]. Curr Drug Targets, 2011, 12(13): 1925-1956.
- [17] Li Y, Tollefsbol TO. Impact on DNA methylation in cancer prevention and therapy by bioactive dietary components [J]. Curr Med Chem, 2010, 17(20): 2141-2151.
- [18] Islam SU, Shehzad A, Sonn JK, et al. PRPF overexpression induces drug resistance through actin cytoskeleton rearrangement and epithelial-mesenchymal transition [J]. Oncotarget, 2017, 8(34): 56659-56671.
- [19] Friedmann E, Hauben E, Maylandt K, et al. SPPL2a and SPPL2b promote intramembrane proteolysis of TNF α in activated dendritic cells to trigger IL-12 production [J]. Nature Cell Biology, 2006, 8(8): 843-848.
- [20] Mentrup T, Loock AC, Fluhrer R, et al. Signal peptide peptidase and SPP-like proteases—possible therapeutic targets? [J]. Biochimica et Biophysica Acta Molecular Cell Research, 2017, 1864(11PtB): 2169-2182.
- [21] Manzo T, Sturmheit T, Basso V, et al. T cells redirected to a minor histocompatibility antigen instruct intratumoral TNF expression and empower adoptive cell therapy for solid tumors [J]. Cancer Research, 2017, 77(3): 658-671.

(责任编辑:张辉洁)